Electron Likelihood in p14

Joseph Kozminski, Robert Kehoe, Harry Weerts
Michigan State University, East Lansing, MI 48824

Su-Jung Park, Arnulf Quadt
Physikalisches Institut, Universität Bonn

John Gardner, Shabnam Jabeen
University of Kansas, Lawrence, KS 66045

DØ Note xxxx

Draft 1.1

November 22, 2003

Abstract

Monte Carlo studies of the electron likelihood are shown. An updated method to discriminate electrons from background processes using a likelihood function is also presented. This likelihood is tuned for p14, has better signal efficiency and fake rejection than the p13 version, and uses the central preshower detector (CPS) for the first time. A new likelihood calculator is also introduced.

1 Introduction

The identification of isolated high-p_T electrons is essential in many analyses, including top quark and electroweak measurements and new phenomena and Higgs searches. Therefore, it is crucial to have a method for efficiently selecting these electrons while rejecting backgrounds that look electron-like.

A likelihood function, which replaced central track confirmation with a likelihood confirmation, was developed by Whiteson and Phaf [1], based on the the likelihood used in Run I[2]. In this note, Monte Carlo studies based on this likelihood are presented. An improved likelihood function, tuned to p14 data, with new variables considered, optimized, and implemented is then discussed. Finally, a new likelihood calculator is introduced.

2 Electron Reconstruction and Preselection

High-p_T electrons are reconstructed from electromagnetic (EM) clusters in the calorimeter. When such a cluster is found, one looks to the tracking detectors for confirmation since an isolated high-p_T track should be in the vicinity of the cluster. The central preshower detector (CPS) can now also be used to confirm the presence of an electron in the central calorimeter (CC) region, $|\eta| < 1.1$. (In the future, the forward preshower detectors (FPS) will likewise be available for confirming electrons in the calorimeter endcaps (EC), $1.5 < |\eta| < 2.5$.)

In the reconstruction, an EM cluster is defined to be a group of towers in a cone of radius $R = \sqrt{\Delta\eta^2 + \Delta\phi^2} = 0.2$ around the highest energy tower. A true EM shower will deposit most of its energy in the first few layers (EM layers) of the calorimeter. Thus, it will have a large EM fraction, $f_{EM} = E_{EM}/E_{tot}$, where E_{EM} is the cluster energy deposited in the EM layers and E_{tot} is the total energy of the cluster.

EM showers should also develop with a logitudinal and lateral shape comparable to that of an electron. Each cluster is assigned a χ^2_{Cal}, or H-Matrix, based on 8 parameters which compare the values of the energy deposited in each layer of the EM calorimeter and the total shower energy with the average distributions obtained in Monte Carlo. Electrons tend to have small H-Matrix values.

Electron candidates also tend to be isolated in the calorimeter. Therefore, the isolation fraction,

$$f_{iso} = \frac{E_{tot}(R < 0.4) - E_{EM}(R < 0.2)}{E_{EM}(R < 0.2)}, \tag{1}$$

tends to be small, meaning that there is not much calorimeter energy in a halo around the EM cluster.

Before the development of the likelihood, $Prob(\chi^2_{EM-trk})$, where

$$\chi^2_{EM-trk} = \left(\frac{\delta\phi}{\sigma_\phi}\right)^2 + \left(\frac{\delta z}{\sigma_z}\right)^2 + \left(\frac{E_T/p_T - 1}{\sigma_{E_T/p_T}}\right)^2, \tag{2}$$

was used to confirm a track match to an electron. (In equation 2, $\delta\phi$ is the difference in ϕ between the extrapolated track impact at the EM3 floor of the calorimeter and the cluster position in the EM3 floor; δz is the difference between the vertex position calculated from the track and that calculated from the EM cluster; E_T/p_T is the ratio of the transverse energy in the calorimeter to the transverse momentum of the track; and σ_ϕ, σ_z, and σ_{E_T/p_T} are the root-mean-squares (RMS) of the experimental distributions of the associated quantities.) The match is better as the probability tends toward 1; nevertheless, $Prob(\chi^2_{EM-trk}) > 0.01$ is a tight track match by definition[3]. This track match, however, is costly in terms of signal efficiency and does not reject background as well as the likelihood[1].

3 Sources of Background

The sources of instrumental electron backgrounds, or "fake" electrons, are believed to be:

- π^0 showers which overlap a track from a nearby charged particle.
- Photons which convert to e^+e^- pairs.
- Charged pions that undergo charge exchange in the detector material.
- Fluctuations of QCD shower shapes.

In Run I, the primary sources of background were identified to be π^0 overlaps and photons conversions[4]. In fact, these backgrounds could be separated using dE/dx, a measurement of energy loss, and transition radiation measured by the transition radiation detector (TRD). Conventional cuts on these quantities were rather inefficienct, leading to the development of the likelihood in Run I. In Run II, there is no TRD, but the improved tracker and the preshower detectors provide other tools which could be used to separate the backgrounds. At this point, however, many of these tools are not yet understood well enough to be fully utilized; hence, these backgrounds are dealt with together for now. Nonetheless, studies of these backgrounds in the Monte Carlo have contributed to a better understanding of the backgrounds as seen below.

In order to distinguish real electrons from fakes, certain characteristics of these fakes must be considered in trying to choose the best discriminating variables. A π^0, for example, is typically produced in association with charged hadrons. Because of this, the calorimeter can be used to pick up signs of hadronic activity around the EM cluster. Moreover, since the π^0 would have to overlap a track from the charged hadrons in order to fake an electron, the track match could be poor; the track would not necessarily be isolated; and E_T/p_T would not tend toward 1, as expected for good electrons.

Photon conversions typically look very electron-like in the calorimeter, though they may be slightly wider than an electron shower. However, one would expect a second track very close to the EM cluster which could be resolved by the tracker or CPS. Also, E_T/p_T would tend to be large. Asymmetric conversions, on the other hand, would be a virtually indistinguishable background since one of the particles would be very soft. Fortunately, asymmetric conversions are very rare.

3.1 QCD Fake Electron Studies

In order to confirm how the QCD background arises, a QCD Monte Carlo sample was examined. There were 104,000 p13 QCD Monte Carlo events with p_T cuts of 20, 40, or 80 GeV. Requiring a loose electron ($f_{EM} > 0.9$, $f_{iso} < 0.15$, H-Matrix8 < 20) with $p_T > 15$ GeV gives 879 loose electron candidates. If, additionally, a tight track match is required, 38 tight electron candidates survive.

Of those 38 tight electron fakes, 19 are faked by a single π^0 or η, 12 by several π^0 or η, three by charged pions, one by a photon that was radiated by a quark and the remaining three are non-isolated actual electrons from meson decays.

If, instead of using the trackmatch, the tight requirement based on the likelihood as described in [1] with a cut at 0.4 is used, only five tight electron candidates survive. Two are faked by a single π^0 or η, two by several π^0s or ηs, and one by a charged pion. See Table 1 for a summary.

Therefore, it can be concluded that the QCD background for electrons is still mainly made up of neutral pions.

Requirements	total	by π^0/η	by x·(π^0/η)	by $\pi^\pm$	by u→ γ...	by meson decays
loose electron ($p_T > 15$ GeV)	879					
loose electron ($p_T > 15$ GeV) + trackmatch	38	19	12	3	1	3
loose electron ($p_T > 15$ GeV) + likelihood + spatial trackmatch	5	2	2	1	0	0

Table 1: Number of fake electrons in QCD.

4 Electron Likelihood in Monte Carlo

4.1 Studies of the Original Likelihood

The likelihood presented in [1] uses 6 parameters:

- $\chi^2_{spatial}$
- E_T/p_T
- H-Matrix8
- f_{EM}
- Distance of Closest Approach (DCA)
- Distance in R to second closest track to the EM cluster.

In order to reproduce the likelihood studies of [1] in Monte Carlo the fake electron statistics in the available QCD Monte Carlo samples is too low. But knowing from the previous study that, in QCD events, electrons are mainly faked by pions, single pion samples can be used as the representative background sample. 84,000 single π^0 events and 173,000 single charged pion events were generated in p14.03. Together they serve as the background sample, mixed in this ratio of 1:2, because, according to strong isospin symmetry, they are produced in that ratio. For the signal sample, 20,100 single electrons were generated as well in p14.03. As can be seen in Figure 1, all samples were generated with flat p_T, η and ϕ distributions and the same amount of negatively and positively charged particles.

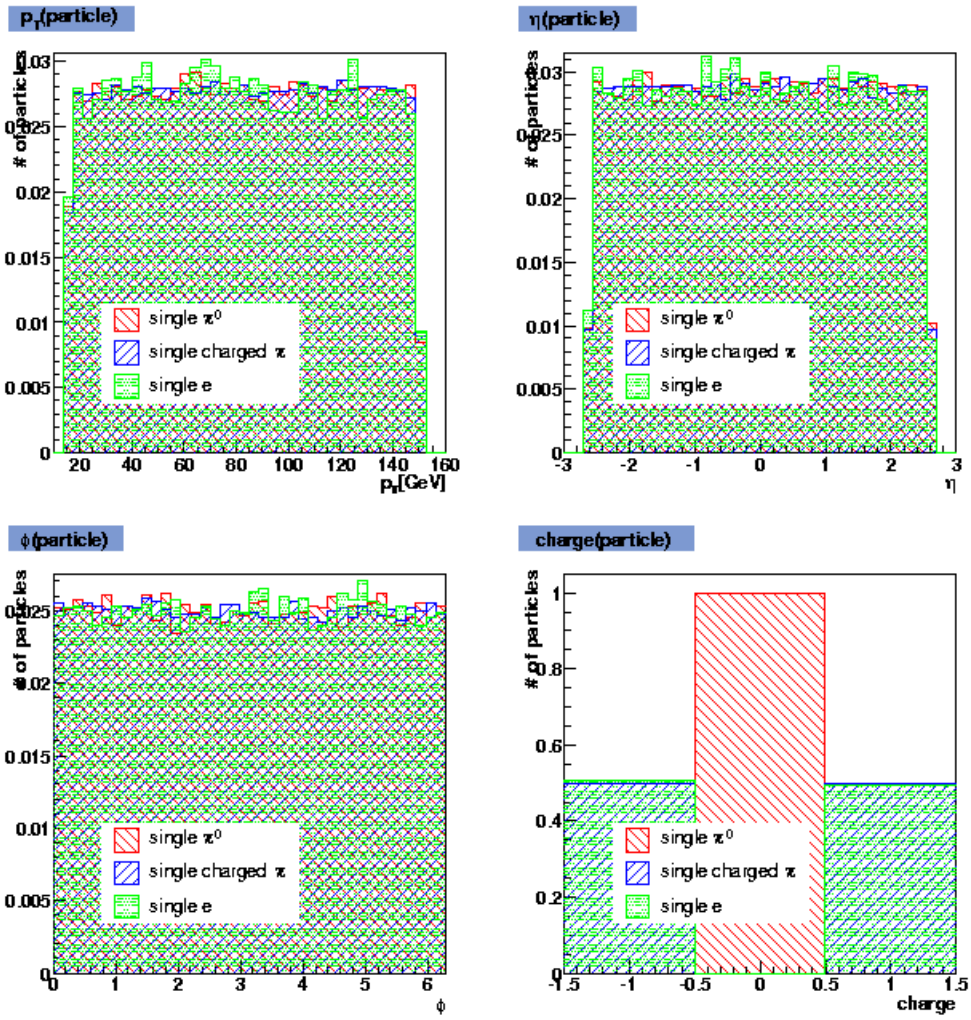

Figure 1: p_T, η, ϕ and charge distribution of generated single particles.

Sample	# of generated events	# of events passing in CC	# of events passing in EC
e	20,100	7,551	6,139
π^0	84,000	3,500	4,101
$\pi^{\pm}$	173,000	198	38

Table 2: Number of single particle events passing the preselection.

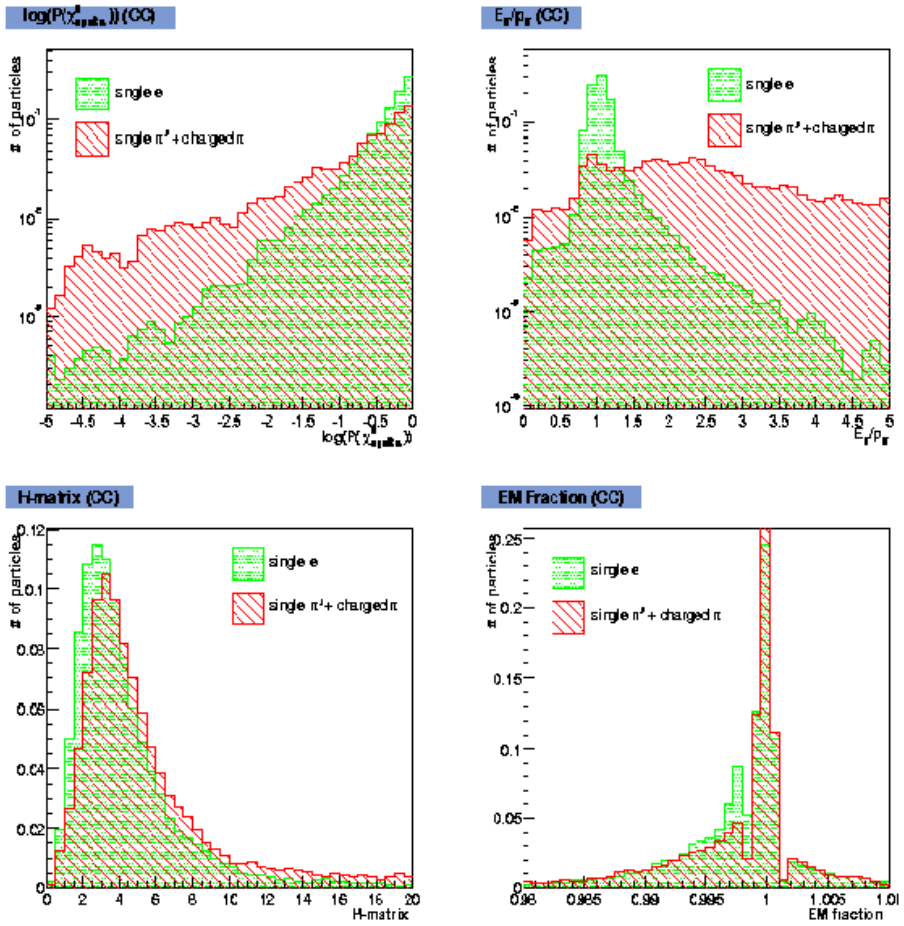

Figure 2: The four likelihood input variables in CC. Top left: logarithm of the χ^2-probability of the spatial trackmatch ($\log(\mathrm{P}(\chi^2_{spatial}))$), top right: E_T of the calorimeter cluster / p_T of the track, bottom left: H-matrix8, bottom right: EM fraction.

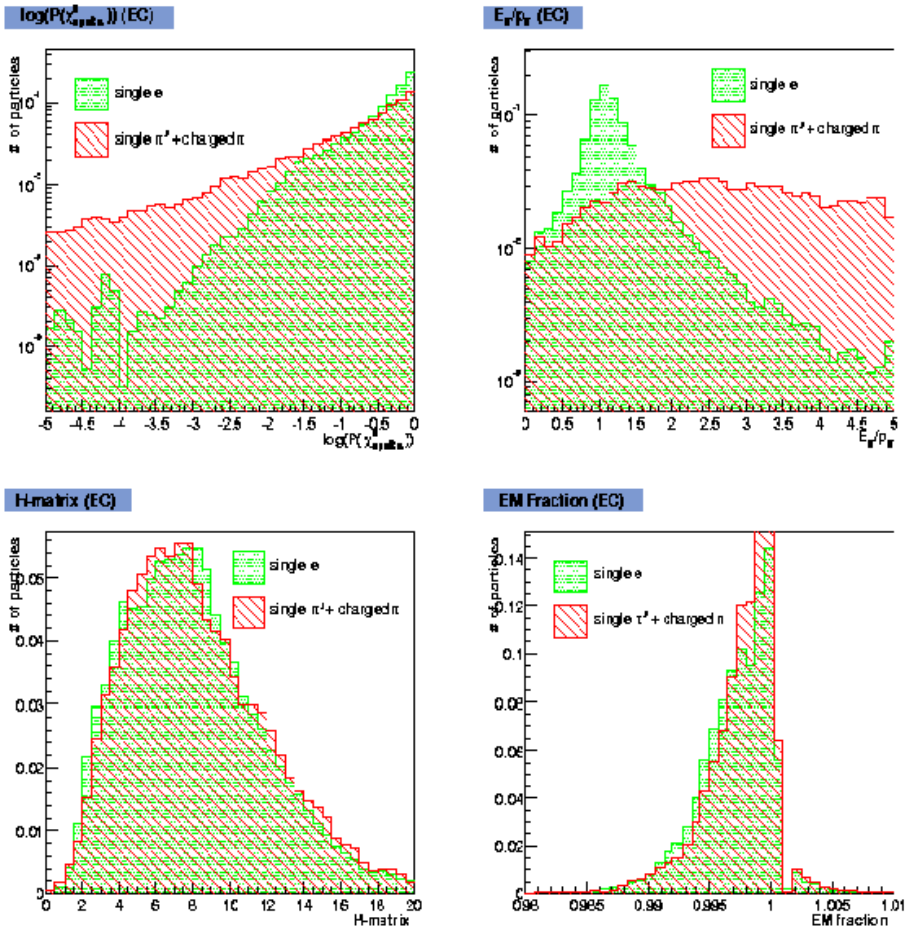

Figure 3: The four likelihood input variables in EC. Top left: logarithm of the χ^2-probability of the spatial trackmatch ($\log(\mathrm{P}(\chi^2_{spatial}))$), top right: E_T of the calorimeter cluster / p_T of the track, bottom left: H-matrix8, bottom right: EM fraction.

The preselection was kept as similar as possible to the one made in data, requiring one EM cluster with $p_T > 20$ GeV, $f_{EM} > 0.9$, $f_{iso} < 0.15$, and H-Matrix8 < 20 with an associated track candidate with $\Delta\phi < 0.05$ and $\Delta\eta < 0.05$. Table 2 shows the number of events passing the preselection for both the CC and EC. It can be seen that fake electrons by π^0s dominate the background.

Of the six input variables used for the likelihood, two do not make much sense in the single particle samples, one being the distance of closest approach (DCA) to the reconstructed primary vertex, because no vertex can be reconstructed with only one track, and the distance to the second closest track, because by construction in most of these events there is only one track. The latter variable should never have been used at all anyway. The distance to the

second closest track, and especially the distribution of this variable, depends on the topology of the event and gives no information about the quality of the electron. An electron from a $t\bar{t} \to e + jets$ event will have a different distribution than an electron from a $Z \to e^+e^-$ event, but still be an electron. A replacement for this variable making use of the track information will be discussed in the next subsection.

The four remaining input variable distributions are shown in Figure 2 for the CC, and in Figure 3 for the EC.

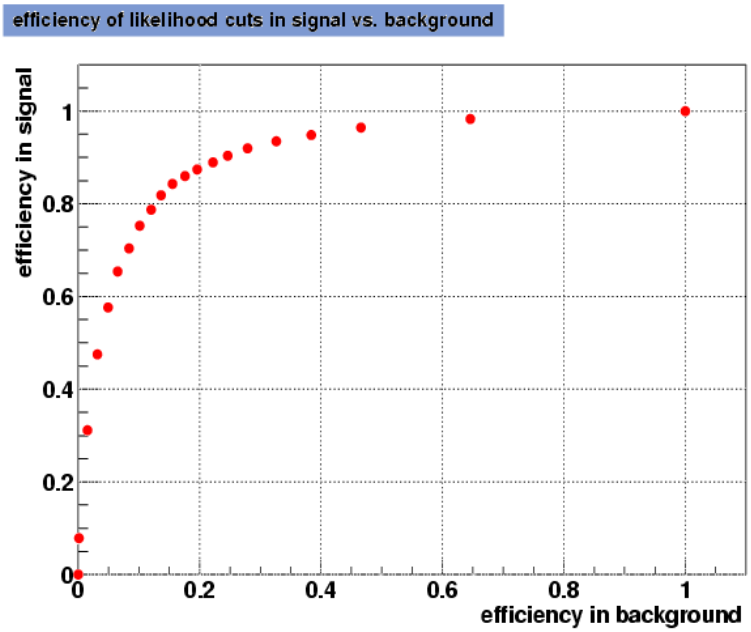

Figure 4: Likelihood cut efficiencies in signal and background with four input variables $\log(\mathrm{P}(\chi^2_{spatial}))$, E_T/p_T, H-matrix and EM fraction.

Calculating the likelihood with those input variables one gets efficiencies in signal and background for different likelihood cuts as shown in Figure 4.

4.2 Number of Tracks in the Vicinity of the Electron

The track information can be utilized to address two different questions. One is identifying photon conversions, which can be done by considering the number of tracks in a small cone around the electron candidate track. The second question is recognizing fake electrons that are part of or very close to a jet by examining the number of tracks in a larger cone around the electron candidate track.

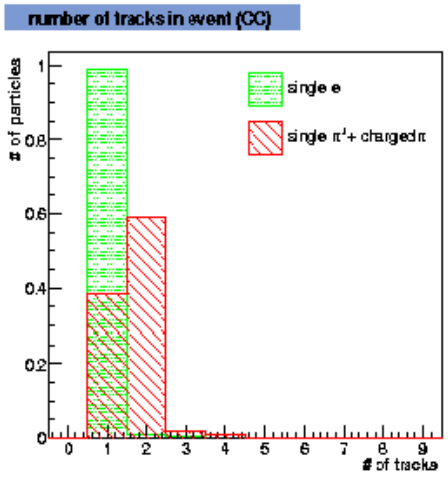

Figure 5: The fifth likelihood input variable in CC, the number of tracks per single particle MC event.

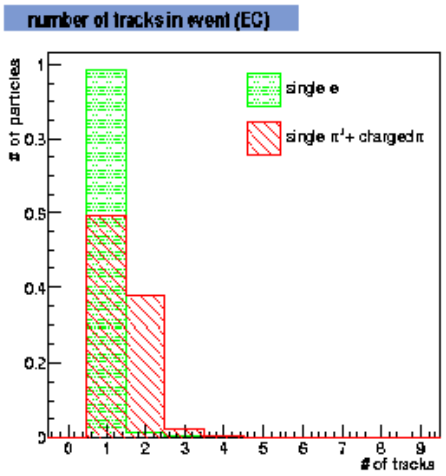

Figure 6: The fifth likelihood input variable in EC, the number of tracks per single particle MC event.

4.2.1 Number of Tracks in a Small Cone

Since it is known that the background is dominated by π^0s, it seems highly desirable to suppress fake electrons by photon conversions. When neutral pions first decay into photons and those convert to e^+e^- pairs, one would expect more than one track; whereas, a real electron should just show one track. In the single particle Monte Carlo samples, it is not necessary to restrict the examination of the number of tracks to a cone because all additional tracks in the event will come from photon conversions. Therefore, the single π^0 sample can instead be used to

determine the optimal size of the small cone, which will be done in the next subsection.

Figures 5 and 6 show the distributions of an additional input variable for the likelihood defined above for these Monte Carlo studies, the number of tracks in a single particle Monte Carlo event. Figure 7 then shows the efficiencies for different likelihood cuts comparing the likelihood with four input variables to the new one with five input variables. The plot shows a significant improvement and so the Monte Carlo studies recommend the introduction of this new variable.

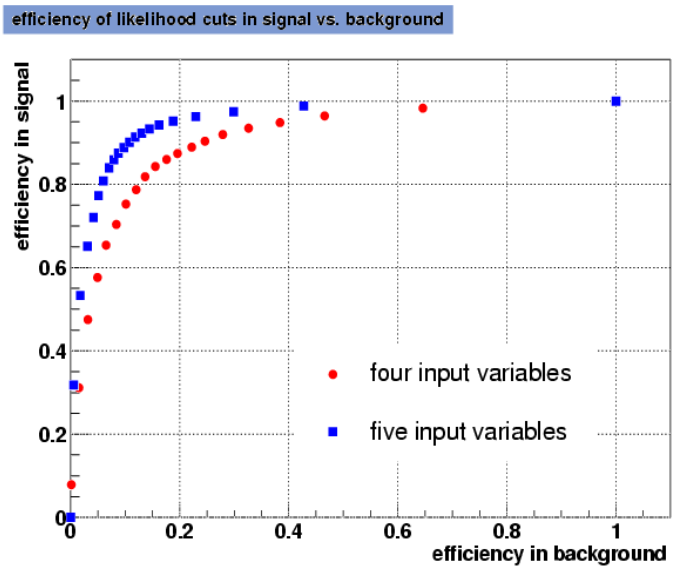

Figure 7: Likelihood cut efficiencies in signal and background.

4.2.2 Size of the Small Cone

In more complicated samples, there are many more tracks in a single event. This makes it necessary to introduce a small cone in which the number of tracks can be counted. The optimal cone size was examined in the single π^0 Monte Carlo sample. Figure 8 shows the distribution of the ΔR distance of the tracks in a single particle Monte Carlo event to the electron candidate track. Additionally, Figure 9 shows the number of tracks collected in a 0.05 cone around the candidate track compared to the number of all tracks in the event, Figure 10 the same for a 0.1 cone. The plots suggest a 0.1 cone to collect all conversion tracks.

4.3 Number of Tracks in a Large Cone

Another track-related variable, which should give a good separation of real isolated electrons to QCD fake electrons, is the number of tracks in a larger cone of 0.4-0.5, the size of a jet.

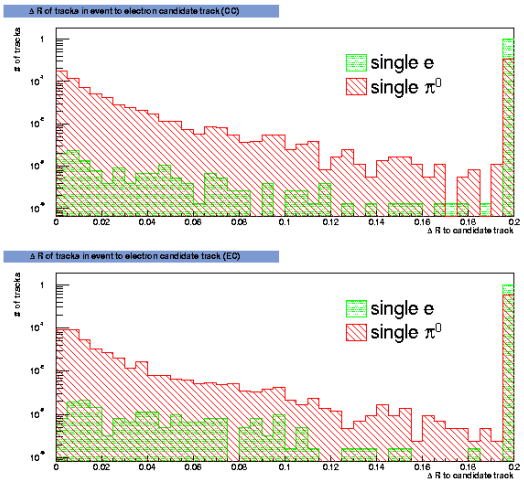

Figure 8: ΔR of the tracks in the event to the electron candidate track in the single electron and single π^0 samples in CC (top) and EC (bottom). If there were no other tracks than the candidate track, or if the track was more than 0.2 away, an entry was made at 0.2.

This should be a good indication if the electron candidate was part of or close to a jet and help suppressing charged pions. However, it cannot be studied in the single particle Monte Carlo samples.

5 Preselection and Likelihood Variables for the p14 Likelihood

5.1 EM Cluster Preselection

Some preselection must be done in order to obtain loose electron candidates on which to apply the likelihood. Electron candidates are required to pass the following calorimeter preselection:

- $f_{EM} > 0.9$.
- $f_{iso} < 0.15$.
- H-Matrix8 < 75.

Note that this H-Matrix cut is a much looser preselection cut than was used with the previous likelihood which required H-Matrix8 < 20. The reason for loosening this cut is to regain some efficiency. The tighter cut has a per electron efficiency of 85% in the CC and 92% in the EC [3].

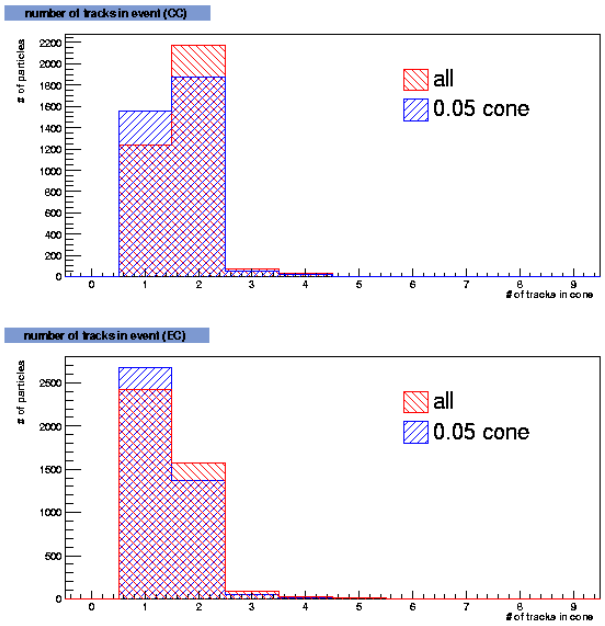

Figure 9: Number of tracks in single π^0 events, comparing the number of all tracks in the event to the number of tracks in a 0.05 cone around the electron candidate track, in the CC (top) and EC (bottom).

The cluster must also have an associated track candidate. An associated track candidate is defined to be a track in a road satisfying

$$|\Delta\phi_{EM,trk}| < 0.05, |\Delta\eta_{EM,trk}| < 0.05.$$

If there is more than one track in this road, the associted track is defined to be the one with the largest $Prob(\chi^2_{EM-trk})$.

5.2 Likelihood Confirmation

A likelihood is a more efficient way of separating good electrons from background than square cuts since a likelihood uses information in the signal and background shapes to distinguish between electrons and background. It also allows variables to be weighted more effectively by their effectiveness in discriminating signal and background compared to conventional cuts.

The new version of the likelihood uses 8 parameters in the CC and 7 in the EC. These variables combine information from the calorimeter, tracking system, and CPS.

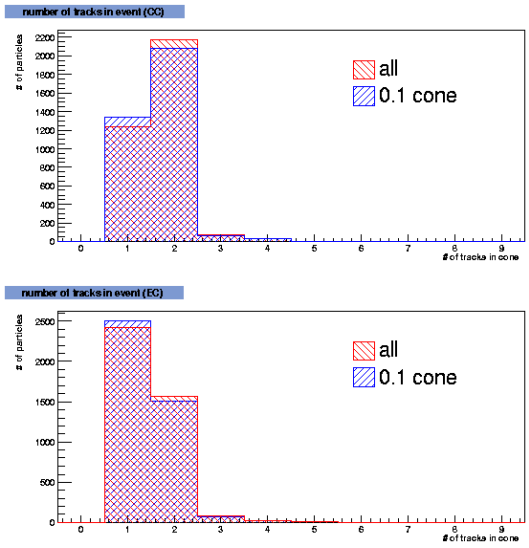

Figure 10: Number of tracks in single π^0 events, comparing the number of all tracks in the event to the number of tracks in a 0.1 cone around the electron candidate track.

5.2.1 Calorimeter Variables

Though there are loose preselection cuts on f_{EM} and χ^2_{Cal}, further information is available in the remaining distributions; thus, these variables are also included in the likelihood.

Adding f_{iso} was also considered; however, when used with the other 8 parameters, this variable gives negligible gain. Therefore, it is not includeded in this version of the likelihood.

5.2.2 Tracking Variables

E_T/p_T is used in the likelihood as it is a very good discriminant for the reasons discussed in Section 3.

Also included is a χ^2 which determines how well a track and an EM cluster are matched in space. This variable is defined as:

$$\chi^2_{spatial} = \left(\frac{\delta\phi}{\sigma_\phi}\right)^2 + \left(\frac{\delta z}{\sigma_z}\right)^2 \tag{3}$$

where $\delta\phi$, δz, σ_ϕ, and σ_z are as defined earlier.

The distance of closest approach (DCA) is again included in the likelihood. This quantity is a measure of the shortest distance of the selected track to the line parallel to the z-axis which passes through the primary vertex position.

Two new tracking quantities are also added to the likelihood. One is the number of tracks in a $R < 0.05$ cone, around and including the candidate track. This parameter is meant to

remove conversions as shown in the Monte Carlo studies. That an 0.05 cone also works well for removing conversions in data is demonstrated in Figure 11.

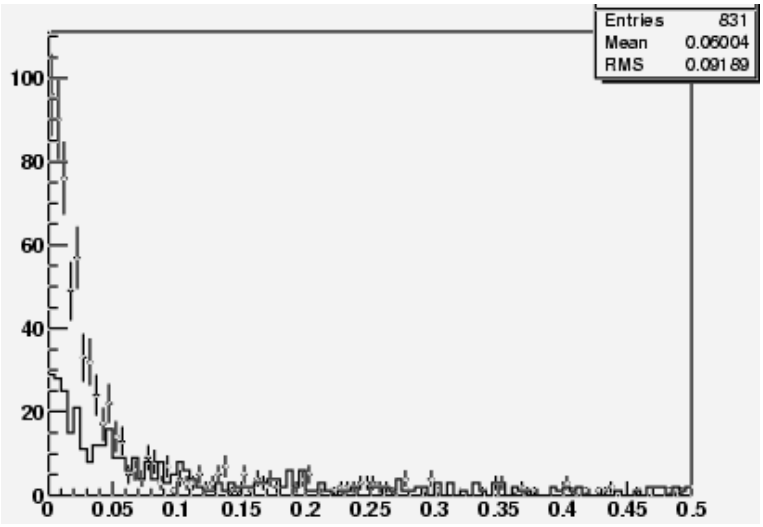

Figure 11: Plot of $|\Delta\phi_{trk}|$ for $1/p_T < 0.5$ [5]. Points with error-bars show opposite-signed pairs while the solid line is the distribution for same-signed pairs. The conversions clearly show up as the bump at $|\Delta\phi_{trk}| < 0.05$.

The other variable is the total track p_T in a $R < 0.4$ cone around, but excluding, the candidate track. Using the number of tracks in a larger cone as suggested by the Monte Carlo studies was tried; however, the sum of the transverse track momenta is a better discriminant. This makes sense since the tracks that are present with good electrons tend to have extremely low p_T (typically well under 0.5 GeV); whereas, tracks from jets tend to have larger p_Ts. This variable is meant to remove π^0s produced in association with charged hadrons. For both of these variables, only tracks whose vertices are within 2 cm of the canditate track's vertex are considered.

Distance in R to second closest track to the EM cluster is omitted for the reasons discussed earlier. However, the new track isolation variable, track p_T in a $R < 0.4$ cone, has just as much discrimination as the previous variable and is not topology dependent.

5.2.3 CPS Variable

In p14, CPS information was included the reconstruction for the first time. This inclusion allows information from a third independent detector to be added to the likelihood in order to enhance discrimination. However, a bug affecting the CPS data is present in reconstruction versions p14.03.01 and p14.03.02. This bug causes all strip energies in the CPS to be set to the saturation value. As a result, energy determination is not possible and position resolution is

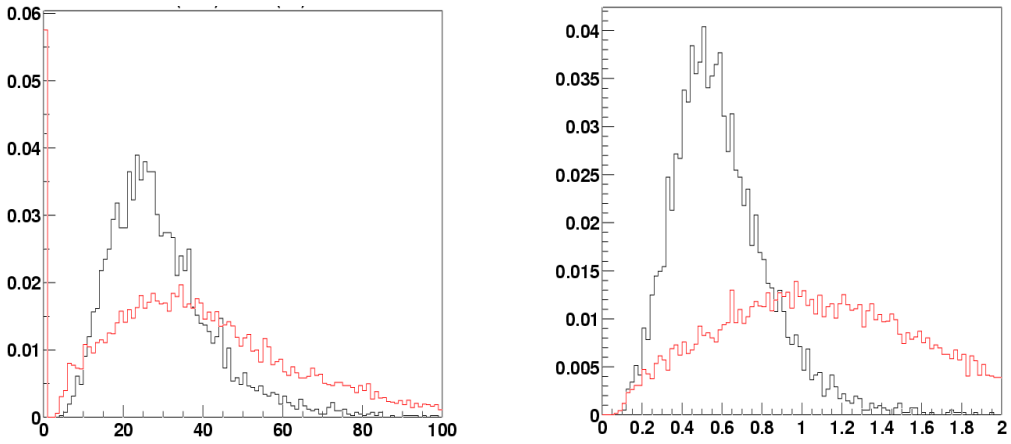

Figure 12: Number of CPS strips (left) and Number of CPS strips / EM Energy (right) for electrons (black) and background (red).

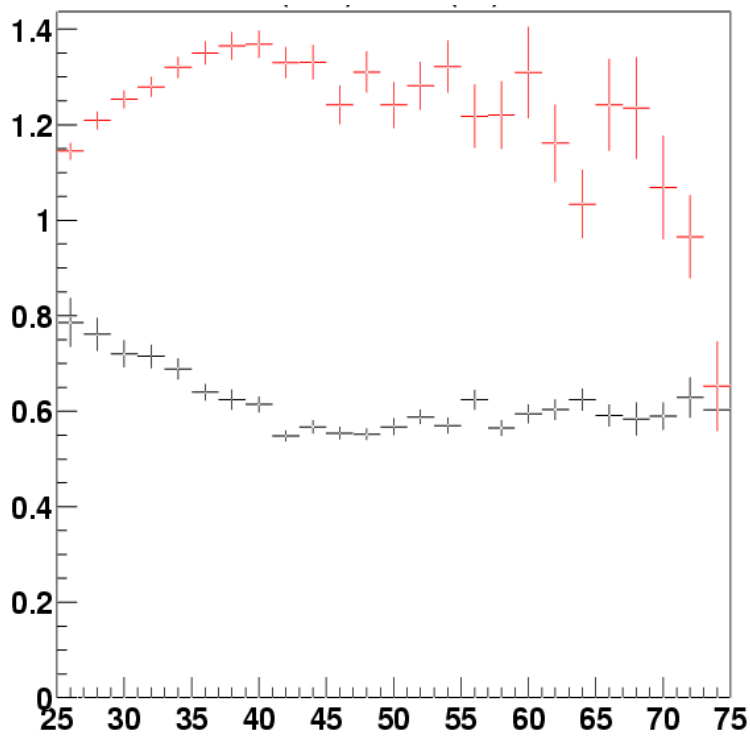

Figure 13: Number of CPS strips/ EM energy in the calorimeter vs. EM energy for electrons (black) and background (red).

worsened. Since a large majority of data was reconstructed with these versions, only variables completely unaffected by this bug, such as number of CPS strips hit, are considered. The most discriminating one is used in the likelihood.

For this CPS variable, the 3D cluster with the largest total number of strips in its single layer clusters is used out of all 3D clusters within a 20 cm radius from the EM object. CPSstripmax is defined as the total number of strips in this 3D cluster divided by the calorimeter EM energy; that is,

$$\text{CPSstripmax} = \frac{\sum_{i=u,v,x} N^i_{hits}}{E_{EM}}, \tag{4}$$

where u, v, and x are the CPS strip layers and E_{EM} is the EM energy. Dividing by the calorimeter energy improves discrimination between electrons and jet fakes (see Figure 12) because, when compared to the fake sample, electrons tend to have higher energy but fewer strips at a given energy. Since number of strips increases with energy, these two effects would tend to cancel each other out if the number of strips were not normalized by the calorimeter energy. Figure 13 shows this variable vs EM energy for electrons and background.

5.2.4 Summary of Likelihood Parameters

To summarize, the parameters used in the likelihood are f_{EM}, H-Matrix8, E_T/p_T, DCA, number of tracks in a $R < 0.05$ cone, total p_T of tracks in a $R < 0.4$ cone around the candidate track, and CPSstripmax (CC only).

6 Likelihood Construction and Performance

6.1 Dataset Definitions

In order to construct the likelihood, a signal and a background sample are needed. The signal sample should be enriched in electrons while the background sample should be enriched in fake electrons. These samples are obtained from p14.03.00, p14.03.01, and p14.03.02 data.

The electron-enriched sample is obtained from $Z \to ee$ events. The selection criteria are:

- Di-electron trigger (2EM_HI, EX_2L20 or EX_2L15_SH15, X=1,2,3) fired.
- 2 EM clusters, each with:
  - $p_T > 20.0$ GeV.
  - $f_{EM} > 0.9$, $f_{iso} < 0.15$, $\chi^2_{Cal} < 75$.
  - Associated track candidate (as defined in the previous section).
- $80 < Mass(e, e) < 100$ GeV.

The background sample is obtained from EM+jet events where the EM object and jet are back-to-back. These events are mainly QCD di-jet and γ+jet events since the W+jet events are removed with a $\not\!\!E_T$ cut. The selection criteria for this sample are:

- Exactly one EM cluster with the above p_T, quality, and track requirements.
- Exactly one good jet with $p_T > 15$ GeV.
- $\not\!\!E_T < 15$ GeV.
- $\Delta R(e, jet) > 2.5$.

6.2 Formulation of the Likelihood

First, distributions for each variable must be obtained from the signal and background samples for CC and EC separately. These distributions are smoothed using linear smoothing techniques and normalized to unit area to produce probability distributions for each variable. These distributions can be found in Figures 14 and 15. Now, these distributions can be used to assign a probability for a given EM object to be signal or background:

$$P_{sig}(\mathbf{x}), P_{bkg}(\mathbf{x})$$

where $\mathbf{x}$ is a vector of likelihood variables. That is, each likelihood variable for the object is given a probability to be signal or background from the probability distributions. Then, assuming no correlations, these individual probabilities are multipled together to give an overall probability for the event:

$$P(\mathbf{x}) = \prod_i P(x_i).$$

Finally, to distinguish electrons from backgroud objects, the following discriminant is used:

$$\mathcal{L}(\mathbf{x}) = \frac{P_{sig}(\mathbf{x})}{P_{sig}(\mathbf{x}) + P_{bkg}(\mathbf{x})}. \tag{5}$$

For electrons, $\mathcal{L}(\mathbf{x})$ tends toward 1; whereas, for background objects, $\mathcal{L}(\mathbf{x})$ tends toward 0.

6.3 Performance

The performance of the likelihood is tested by running over signal and background samples. Figure 16 shows that the likelihood separates signal from background very well after the preselection cuts. This separation power can be seen in Figures 17 and 18 by looking at signal and background efficiencies for different thresholds of the likelihoods shown in Figure 16. These plots clearly show the large relative efficiency gain obtained by starting with a looser H-matrix

Threshold	Signal Eff.	Background Eff.
	Central Calorimeter	
$\mathcal{L}_8$ >0.3	0.939±0.004	0.246±0.003
$\mathcal{L}_8$ >0.4	0.931±0.004	0.218±0.003
$\mathcal{L}_8$ >0.5	0.923±0.004	0.192±0.003
$\mathcal{L}_8$ >0.6	0.910±0.004	0.170±0.003
$\mathcal{L}_8$ >0.7	0.894±0.005	0.149±0.003
$\mathcal{L}_8$ >0.8	0.869±0.005	0.123±0.003
	Endcap Calorimeter	
$\mathcal{L}_7$ >0.3	0.878±0.013	0.226±0.007
$\mathcal{L}_7$ >0.4	0.867±0.013	0.197±0.007
$\mathcal{L}_7$ >0.5	0.851±0.014	0.173±0.007
$\mathcal{L}_7$ >0.6	0.835±0.015	0.154±0.006
$\mathcal{L}_7$ >0.7	0.824±0.015	0.134±0.006
$\mathcal{L}_7$ >0.8	0.799±0.016	0.106±0.005

Table 3: Efficiencies for signal and background at different threshold levels of the likelihood.

preselection. In addition, Figure 17 shows that the new likelihood performs considerably better than the likelihood used in p13 where a good electron was defined to be an object with a likelihood greater than 0.4.

Table 3 shows efficiencies for signal and background (after pre-selection) at different threshold levels of the likelihood in the CC and EC. These values are read off Figures 17 and 18.

7 Likelihood Calculator

A new likelihood calculator, which is included in the *emreco* package, has been written. This calculator is meant to be a stand alone piece of code which anyone can run. The probability distributions, which are the look-up tables, are hard-coded as arrays; hence, there is no dependence on *ROOT* in this version of the likelihood calculator. (When the RCP package is someday made independent of other DØ libraries, these tables can easily be implemented as RCP files if desired.) Moreover, this code should be usable within any analysis framework since the user supplies the likelihood variables, and the calculator simply returns the value of the likelihood.

After the class *LhoodCalc* is instantiated, the member function

double calc_lh(*double* deta, *double* emf, *double* hmx, *double* etpt, *double* dca,

double spchi, *double* ndr05, *double* tpt, *double* cps, *int* cps_flag)

can be used to calculate the likelihood. The input variables are:

- deta: Detector eta of the electron.
- emf: f_{EM}.
- hmx: χ^2_{Cal}.
- etpt: Calorimeter E_T / track p_T.
- dca: DCA.
- spchi: $Prob(\chi^2_{spatial}, N_{dof})$
- ndr05: Number of tracks in an 0.05 cone around (and including) the candidate track.
- tpt: Total p_T of tracks in an 0.4 cone around the candidate track, but excluding the candidate track p_T.
- cps: Number of CPS strips in cluster / EM energy.
- cps_flag: 0 means that the CPS is *NOT* included in the likelihood calculation. 1 means that the CPS is used in calculating the likelihood.

When a likelihood can be calculated, this function returns a likelihood value between 0 and 1 where the event is more signal-like as this value tends toward 1. If the electron is in the inter-cryostat region (ICR), $1.1 < |\eta| < 1.5$, or is in the far-forward region, $|\eta| > 2.5$, the likelihood is not defined and -1 is returned.

8 Further Studies

Further Monte Carlo studies for the improvement of the electron likelihood have been made and presented in [6] and will be described in detail in [7]. These include:

- Comparison of the two-class-likelihood described to a three-class-likelihood, separating the background class into two classes of π^0 background and $\pi^\pm$ background. The three-class-likelihood shows a major improvement for suppressing the charged pion background, but performs a little bit worse on the predominant π^0 background.
- Likelihood with p_T-dependant input variable distributions.
- η-Depandance of the input variable distributions. Although the distributions are in fact η-dependant, using this will not give any additional separation, because signal and background distributions change in the same way.

- Two-dimensional likelihood, calculating separate likelihoods for the π^0 background and the $\pi^{\pm}$ background.

In data, one of the next things to start looking at should be correlations between variables. Since adding new variables no longer seems to add much discrimination power to the likelihood, it is hoped that information in the correlations could be used to further discriminate between signal and background.

9 Acknowledgements

We would like to thank Volker Büscher and Marumi Kado for their help in our studies and Drew Alton, Sean Mattingly, and Serban Protopopescu for showing us how to generate single particle Monte Carlo samples. We would also like to thank Drew Alton for the great help he provided in getting CPS information into the likelihood. Finally, we would like to thank Yuri Gerstein for his suggestions for new variables to look at (namely track p_T in a large cone) and for some very useful discussions.

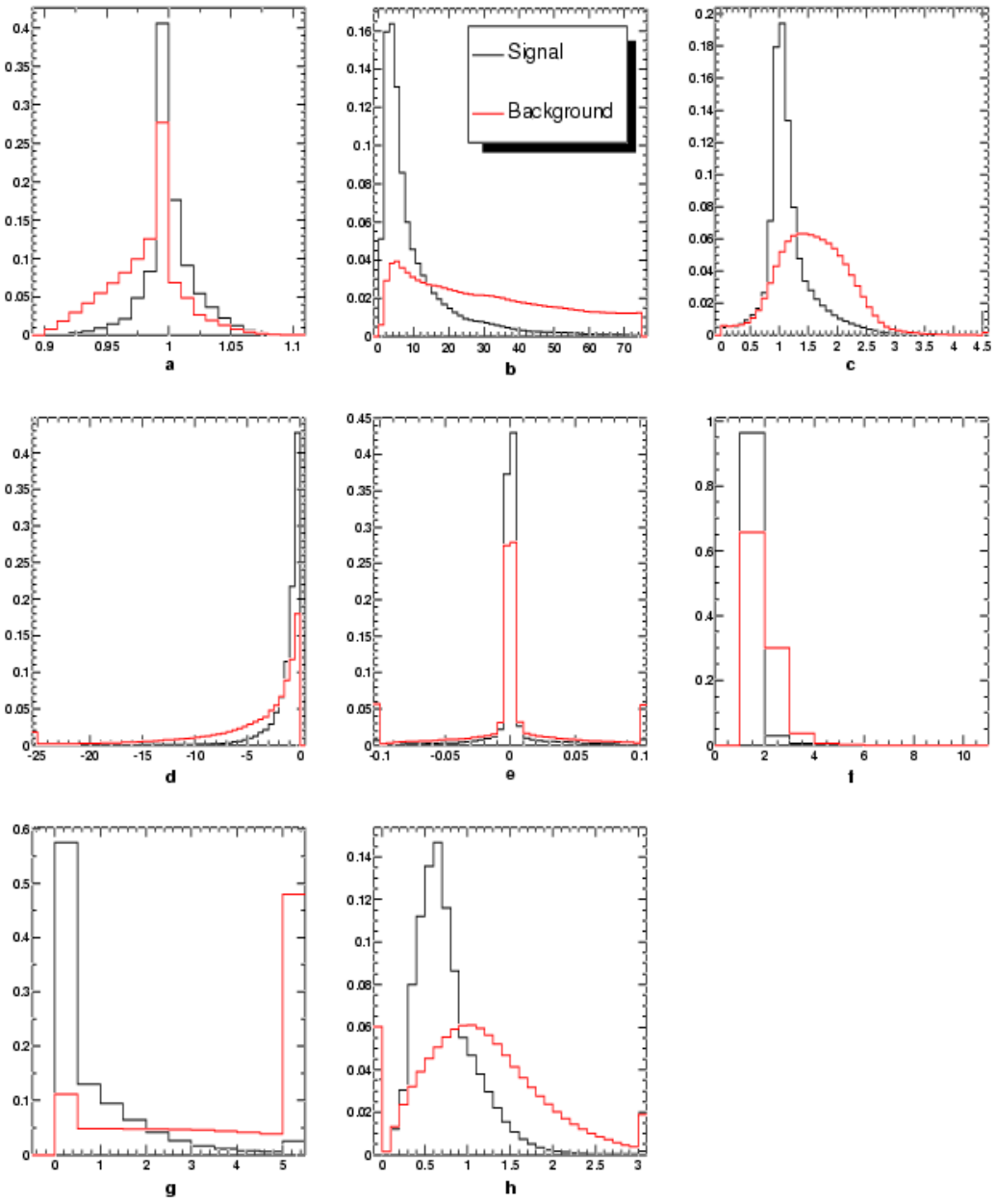

Figure 14: Probability distributions for signal and background in the CC. These distributions are (a)f_{em}, (b)H-Matrix, (c)E_T/p_T, (d)$log(P(\chi^2_{spatial}))$, (e)DCA, (f)Number of tracks in 0.05 cone, (g)Total p_T of tracks in 0.4 cone around candidate track, and (h)Number of CPS strips/ Calorimeter E.

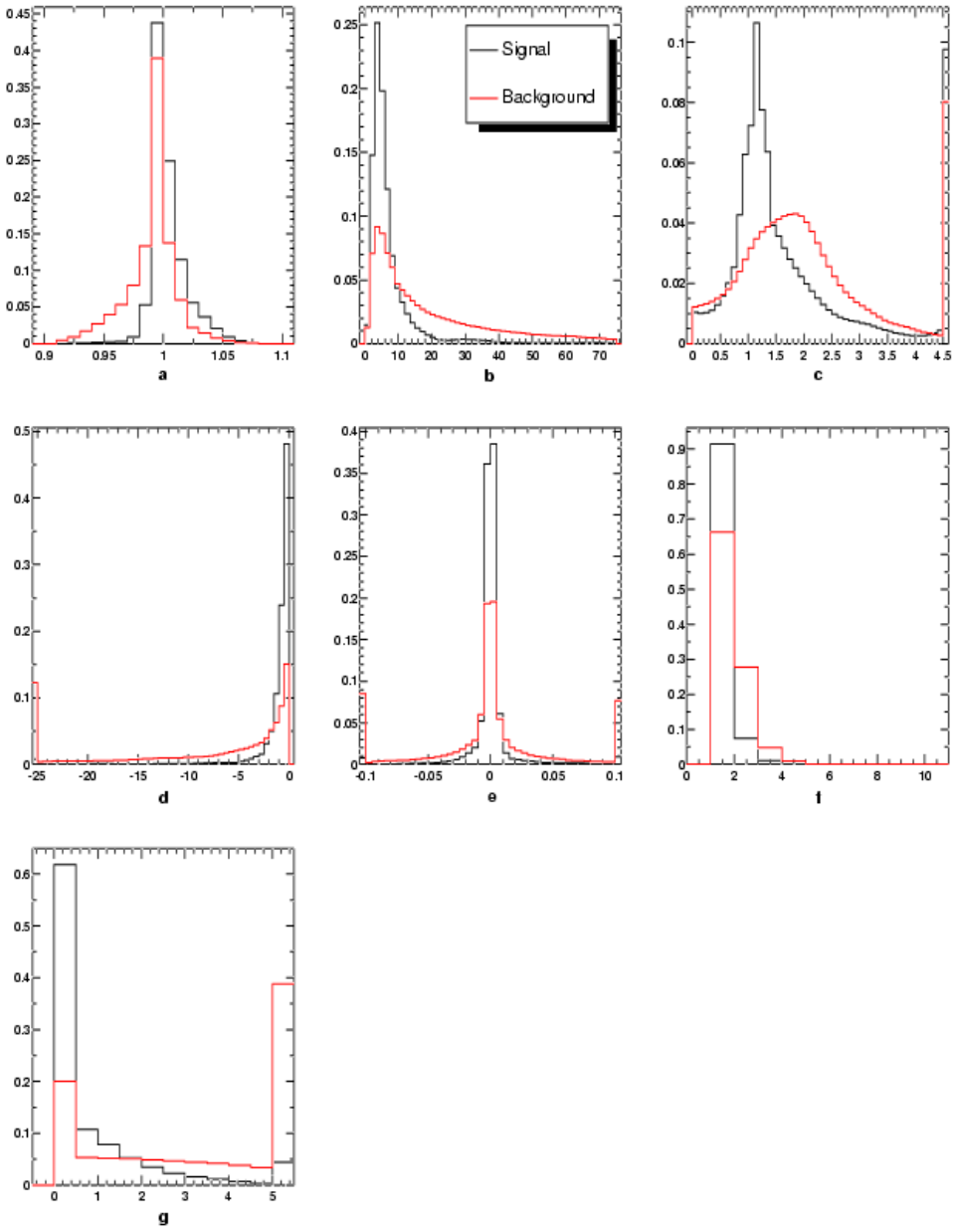

Figure 15: Probability distributions for signal and background in the EC. These distributions are (a)f_{em}, (b)H-Matrix, (c)E_T/p_T, (d)$log(P(\chi^2_{spatial}))$, (e)DCA, (f)Number of tracks in 0.05 cone, (g)Total p_T of tracks in 0.4 cone around candidate track, and (h)Number of CPS strips/ Calorimeter E.

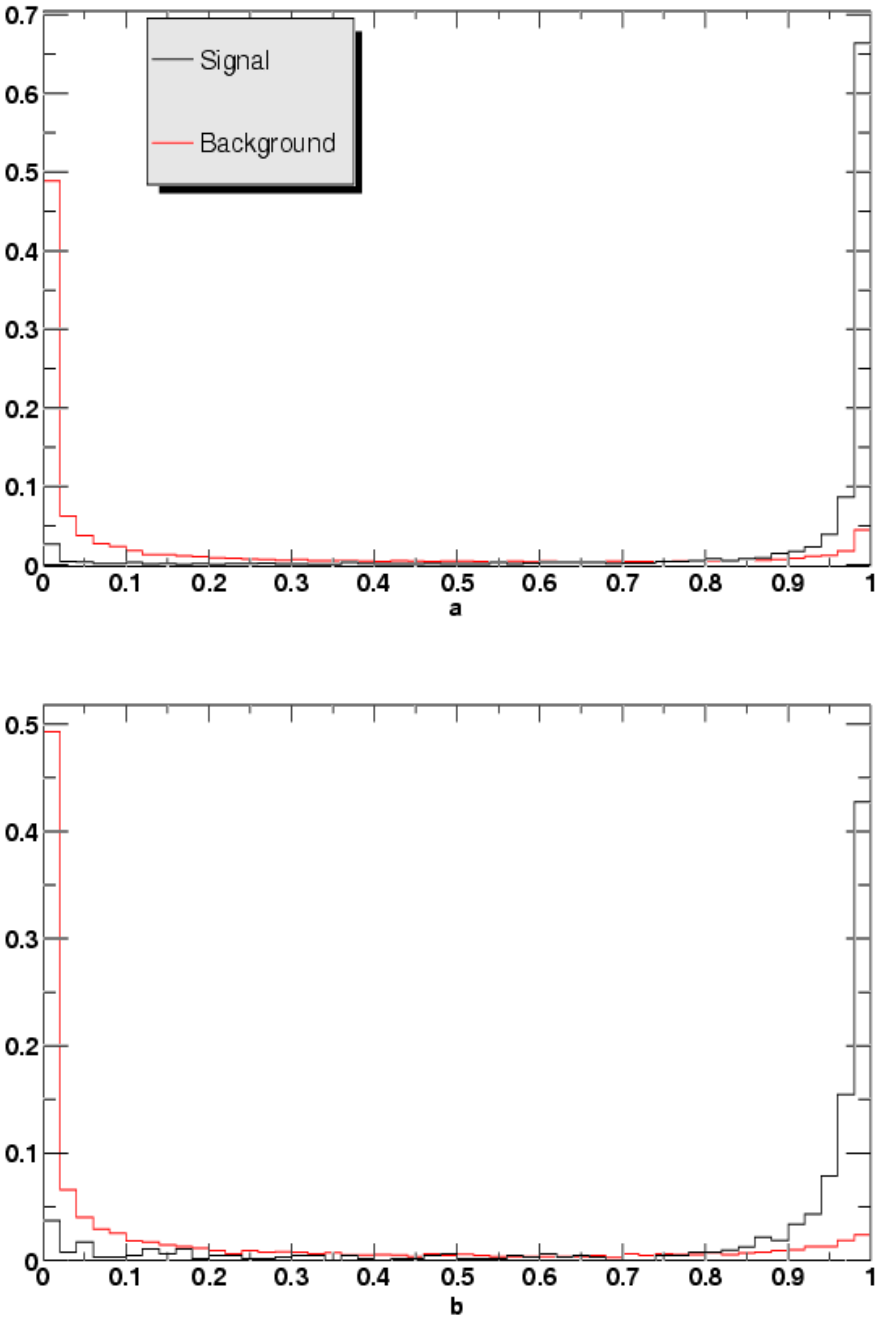

Figure 16: Distributions of the likelihood in the (a)CC and (b)EC.

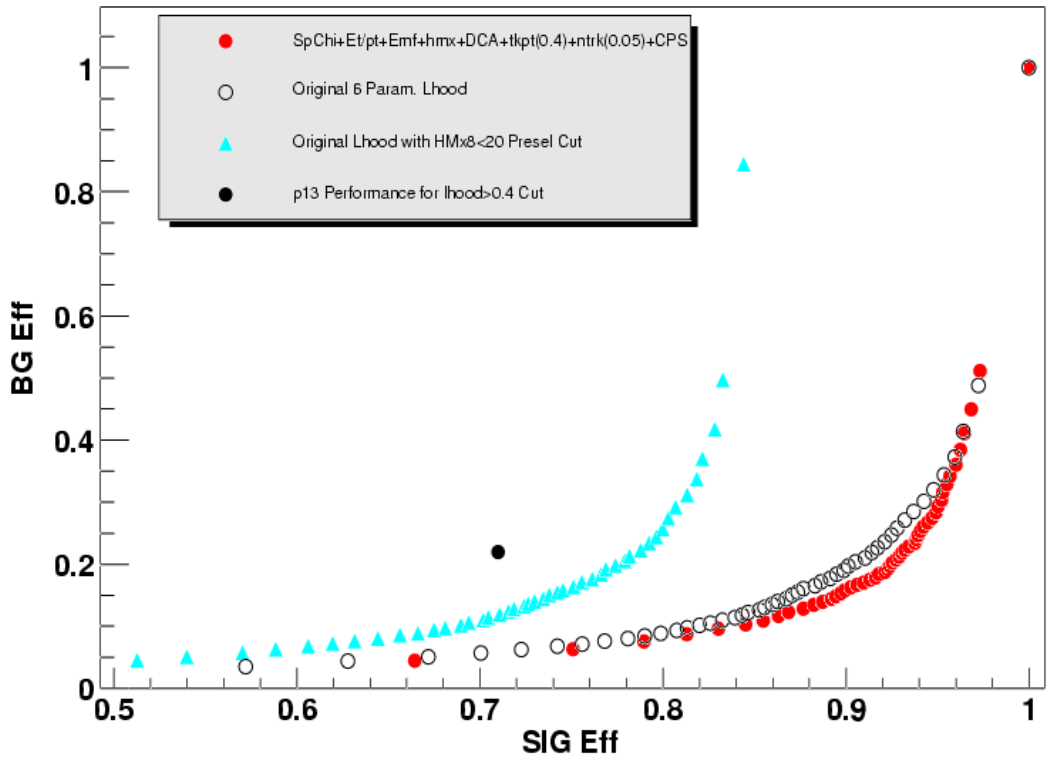

Figure 17: Signal and background efficiencies in the CC for varying thersholds of the CC Likelihood in Figure 16.

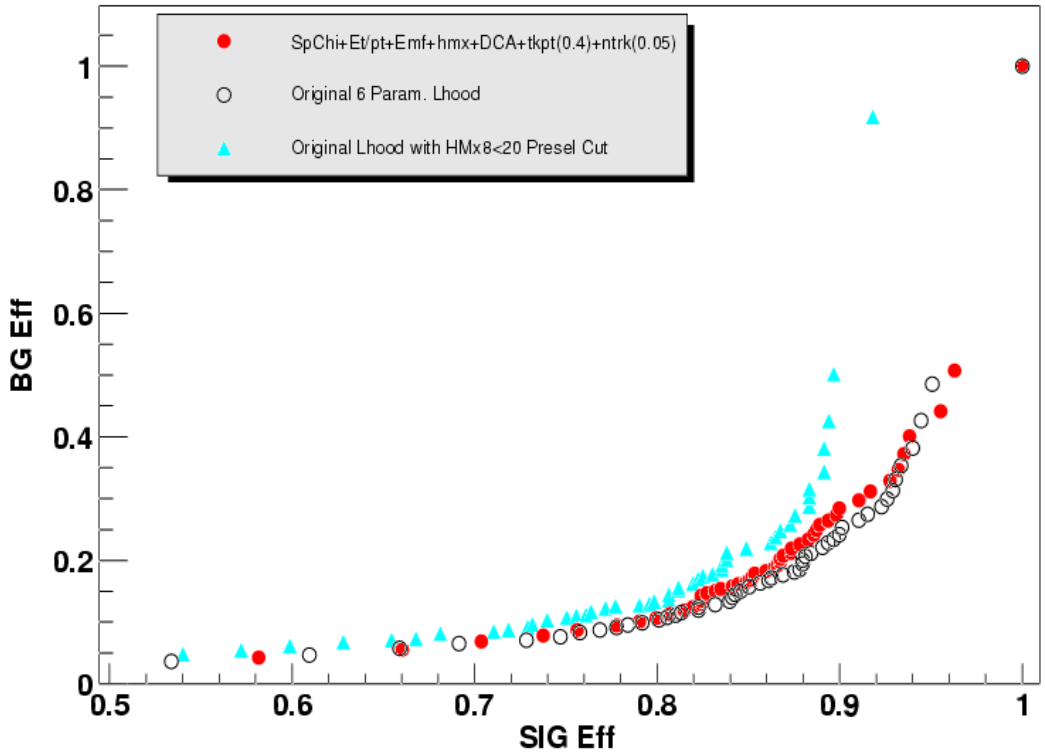

Figure 18: Signal and background efficiencies in the EC for varying thersholds of the EC Likelihood in Figure 16.

References

[1] D. Whiteson and L. Phaf, DØ-Note 4184.

[2] M. Narain and U. Heintz, DØ-Note 2386.

[3] EM ID Certification, v4.1,
http://wwwd0.fnal.gov/phys_id/emid/d0_private/certification/main_v4_1.html.

[4] M. Narain and U. Heintz, DØ-Note 2355.

[5] Y. Gerstein, EM ID presentation,
http://www-d0.fnal.gov/phys_id/emid/d0_private/minutes/20030730yuri.ppt

[6] S.-J. Park, Talk on the *Electron Likelihood in Monte Carlo* at the DØ Calorimeter Workshop on October 6, 2003.

[7] S.-J. Park, Diploma Thesis on the $t\bar{t}$ *cross section in the e+jets channel in Monte Carlo* (in preparation).